干旱区地理
ARID LAND GEOGRAPHY

祁连山及周边降水分布聚类检验和典型流域增雨效果评价

郭小芹[1,2]， 李光明[2]， 孙占峰[2]， 王兴涛[2]
（1. 中国气象局云雾物理环境重点开放实验室，北京 100081；2. 武威市气象局，甘肃 武威 733000）

摘 要：祁连山及周边区域地形复杂，降水分布极不均匀，如何科学地分区分型，对把握研究区域降水分布特征具有极为重要的意义。利用该区域31个国家基本站1961—2020年5—9月降水量资料，采用主成分分析方法（PCA）对该区域降水量进行分析，再通过聚类分析（CAST）对该结果进行显著性检验，最后将分区分型结果应用于人工增雨作业效果评估。结果表明：(1) 祁连山及周边可分成7个区域（Z1~Z7），累积方差贡献率超过78%。(2) 这7个区域分别以乐都、海晏、野牛沟、武威、高台、临泽、刚察为中心点，划分结果与降水量、地理地形、海拔高度显著关联。(3) 依托区域历史回归统计方法对人工增雨作业效果进行评价，发现1992—2020年石羊河流域5—9月绝对增雨量、相对增雨率分别为8.91 mm、6.51%，其中7月最高（6.30 mm、21.86%），8月次之（5.44 mm、16.11%）。基于地面降水量的作业效果评价往往受对比区选择的影响，聚类检验方法不仅有助于复杂地形下降水量的分区研究，还为科学选择对比区提供了客观依据。

关 键 词：聚类检验；主成分分析；分区分型；增雨效果评价
文章编号：

祁连山地处青藏、蒙新、黄土高原交汇地带，位居我国干旱半干旱区内陆腹地，横跨青海省东北部与甘肃省西部地区。整个祁连山区属褶皱断块，自西北向东南依次为大雪山、托来山、托来南山、野马南山、疏勒南山、党河南山、土尔根达坂山、柴达木山和宗务隆山，平均海拔4000 m以上，山脉东西长1000 km，南北宽300 km，范围广大，地形复杂，降水区域性分布特征尤为明显。作为西北干旱区的一条湿舌，祁连山不仅是诸多河流与水系的发源地，也是周边区域生存与发展的屏障。祁连山北侧是石羊河、黑河、疏勒河三大内陆河径流区，是河西绿洲存活的基础；祁连山南侧是大通河、湟水河流经地，是河湟流域生存的命脉。祁连山的存在影响着周边大范围地区，形成一个相互影响且不可分割的整体。近年来祁连山雪线上升、冰川面积锐减、水资源严重短缺等问题日益凸显，成为社会经济与生态环境可持续发展的制约因素，也成为学术界探究的热门话题[1-4]。随着气候变暖，极端天气影响频繁出现，探讨祁连山及周边降水分布与变化规律有着重要的现实意义。

人工增雨是开发祁连山空中云水资源以及缓解水资源短缺的有效手段之一，其有效性与科学性已经被越来越多的研究者所重视[5-7]。作为人工增雨试验研究的技术难点，作业效果评价是一项极其复杂而重要的工作。由于自然降水变率极大，评价对象本身存在一定的不确定性，各种复合因子相互影响，要准确客观地厘清界限是比较难的。比如，如何将地面降水量中人工增加部分从自然降水量中识别出来，如何剔除“假效果”而实现精准评价以及如何建立概念模型等问题[8-9]。区域历史回归统计检验是增雨效果评价的基本方法，但往往受分区分型方法与影响区、对比区设计的左右，很显然，不同对比区将会生成不同的评价效果。无论是统计检验、物理检验和数值模拟检验，在对比区选择上，准确性是很难把握的。要科学评价增雨效果，精准分区尤为重要。聚类检验为对比区选取提供了一

收稿日期：2021-07-30； **修订日期**：2021-11-14
基金项目：第二次青藏高原综合科学考察项目（2019QZKK0104）；国家重点研发计划（2019YFC1510302）；甘肃省气象局科研项目（Ms 2022-21）资助

作者简介：郭小芹（1965-），女，高级工程师，主要从事应用气象与人影应用研究. E-mail: gxq9179@126.com

种量化方法，为增雨效果评价的定量、客观提供了科学的解读依据。

多年来气候区划边界问题因受制于人为因素左右，其地理分类方法备受质疑，而主成分分析(Principal component analysis，PCA)在一定程度上规避了这种缺陷，但在荷载值上依然存在临界值勘定问题，因而在一定意义上尚不能实现科学准确的数值分类。聚类分析又称群分析，是分类学中依据事件的某些相似性特征进行归类的一种多元方法，在分析中多选用距离系数、相似系数、相关系数、相关距离系数等指标以实现相似性度量[10-11]。尽管如此，很多学者依然认为该方法缺乏统计学基础。幺枕生[12]提出的聚类统计检验方法(Cluster analysis with statistical test，CAST)则有效弥补了分区分类过程中显著性检验的重大缺陷，尤其在中心聚类评价上具有广泛的适用价值[13-14]。

根据聚类统计检验分析的基本思路，本文对祁连山及周边降水空间分区分型的合理性进行了深入探讨，研究结果将有利于剖析复杂地形下不同降水区域的空间变化特征，为准确把握祁连山降水变化特征与分布规律提供科学依据。

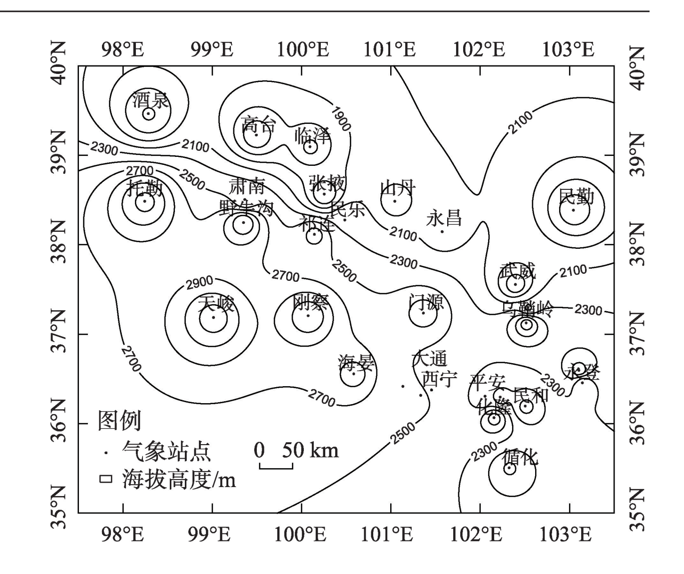


图1 祁连山及周边海拔高度变化与气象站点分布

Fig. 1 Changes of altitude and spatial distribution of meteorological stations in Qilian Mountains and its surrounding areas

1 数据与方法

1.1 数据来源

本文数据来源于中国气象局国家气象信息中心提供的祁连山以及周边区域31个国家基本站1961—2020年5—9月降水量资料，研究所用序列经过严格质量控制。如遇个别时段缺测，则按照维尔达(WILD)综合订正法取邻近站点展开序列订正，在通过适当性判据之后，实现该站点缺失值的插补[15]。祁连山及周边海拔高度变化与气象站点分布见图1。祁连山西高东低，最大海拔高度差2084 m，为了消除地理因素引起的平均值以及变率的影响，在分析前对上述资料进行了标准化处理，形成祁连山降水空间场数据集。分析过程采用多种统计方法并借助SPSS、Matlab编程以及ArcGIS等软件实现了降水特征的分区分型研究。

1.2 研究方法

1.2.1 主成分分析 特征值是衡量主成分影响度的重要指标，代表了引入该主成分以后能够解释源数据信息量的多少，一般以特征值≥1作为考量主成分重要性的标准。方差贡献率(或累积贡献率)用于判断主成分的份量，其数值越大，表明该主成分(或上述主成分)包含的源数据信息量就越大。荷载值反映了主成分与各个变量间的密切程度，在标准化变量下，任意一个主成分的荷载值实质上就是主成分与原变量的相关性指标[16-17]，也就是说，在一定意义上荷载值就类似于相关系数，其值越大，表明该因子对当前变量的影响程度越大。

主成分分析[18-20]可以归纳为以下几个步骤：(1)对降水量数据进行标准化处理，以消除地域因素带来的系统性误差；(2)计算标准化样本的相关系数矩阵；(3)按照特征值和方差贡献率的大小提取因子数量，取得阈值范围内的主分量；(4)对各个主分量进行分析，考察可解释性，如有必要可对因子进行旋转，通过荷载值以寻求更好的解释方式。

1.2.2 聚类检验 聚类分析可将样本划分为不同类群，通过评价分析，其结果具有综合性、客观性和科学性，近年来在各个领域得到广泛应用[21-23]。气候分型区划理论研究已经证明了统计聚类检验与主成分分析的关联性，并认为在一定程度上聚类检验是主成分应用于分区分型的理论基础，并通过多种方法证实了这种关联的有效性[16-17]。聚类分析常用距离来度量数据间的亲疏关系，诸如明可夫斯基距离、曼哈顿距离、欧几里得距离、契比雪夫距离等。将变量间这种距离用C_{ij}表示，C_{ij}与相关系数具有一致性，同时又具备正态分布性，通过χ^2分布函数的

构造，可以实现聚类过程的显著性检验。

假设某气象要素具有时间序列长度为n、测站数量为p所构成的变量场X，已经证明，测站X_i和X_j之间的距离可以用相关系数(r_{ij})表示。根据幺枕生[12]理论，定义自由度$\eta-1$的分布统计量：

$$\chi^2 = 2n\sum_{i<j}\omega_{ij}(c_{ij}-\hat{c})^2 \quad (1)$$

$$c_{ij} = \sqrt{1-r_{ij}} \quad (2)$$

$$\omega_{ij} = 1/(1-r_{ij}) \quad (3)$$

$$\hat{c} = \sum_{i<j}\omega_{ij}c_{ij}\Big/\sum_{i<j}\omega_{ij} \quad (4)$$

式中：$\eta=\frac{1}{2}q(q-1)$，是与自由度相关的统计量；q为同一类型的变量个数，对于气象变量场X来说，q就是变量场中同一分区分型结果中的测站数量($q\leqslant p$)，其中i=1, 2, …, $q-1$，j=1, 2, …, q；c_{ij}表示距离系数；ω_{ij}表示距离权重；$\hat{c}$表示同一类型中的极大似然估计值。

根据式(1)~(4)求出相应站点χ^2统计量，在给定信度α下，以$\chi^2 \geqslant \chi^2_{1-\alpha}(\eta-1)$为临界值，若$\chi^2 < \chi^2_{1-\alpha}(\eta-1)$，则表明该测站可归属于同一区域，取该区域最大荷载值为聚类中心，所有符合条件站点即为同一类别的范围。

1.2.3 *区域历史回归检验* 区域历史回归检验方法是基于地面降水量的效果评价方法，该方法立足对比区、影响区地面降水量相关分析，采用最小二乘法建立一元线性回归方程，预测出影响区的自然降水量，通过自然降水量估计值与实测降水量的分析比较，得出绝对增雨量和相对增雨率。该方法在人工增雨作业效果检验技术指南中已有介绍，详见文献[24]，这里不再赘述。

2 结果与分析

2.1 降水量分布基本特征

祁连山及周边地区降水量分布与祁连山脉走向基本一致。以5—9月为例，东段降水在350 mm以上，西段降水只有100 mm左右，表现出极不均匀性。降水量年度分布不均，以1961—2020年为例，全域多年平均降水量343.7 mm，降水量最大值441.8 mm(1967年)，降水量最小值240.2 mm(1991年)，降水变率±30%，高值区多在祁连山区。降水季节性特征明显，夏季集中，冬季稀少，5—9月降水量占全年85%以上。降水地域分布极不均匀，以2016年为例，张掖降水量还不足刚察的1/45，出现了50 a不遇的气象干旱。研究发现，祁连山及周边地区降水量与地理位置相关联，降水量与纬度呈显著负相关，与海拔呈显著正相关，也就是说纬度越高，降水量越少，海拔越高，降水量越多，这种规律使得祁连山及周边地区降水呈现出东多西少、南多北少的特征，最大降水带分布在2500~3500 m区域，也说明西南暖湿气流与夏季风对该区域降水形成有着极其重要的影响。

2.2 降水量的聚类分析

建立1961—2020年祁连山及周边31个测站5—9月地面降水量原始矩阵，对标准化后的相关矩阵做主成分分析，然后以各个主成分对应的高荷载值作为聚类中心，依据式(1)~(4)给出一定信度下的分区检验判据，对分区结果进行分析评价。

首先对降水量相关矩阵进行主成分分析(表1)，依据特征值以及方差贡献率筛选之分量。从表1可以看出，旋转前后7个主分量其累积方差贡献率均为78.135%，经过旋转之后主分量C_3~C_7的方差贡

表1 祁连山及周边降水量的主成分分析结果

Tab. 1 Results of principal component analysis of precipitation in Qilian Mountains and its surrounding areas

主分量	主成分分析(PCA)			鲁棒主成分分析(RPCA)		
	特征值	方差贡献率/%	累积贡献率/%	特征值	方差贡献率/%	累积贡献率/%
C_1	10.842	34.973	34.973	6.694	21.592	21.592
C_2	4.547	14.666	49.639	4.182	13.491	35.083
C_3	2.942	9.490	59.129	3.824	12.334	47.417
C_4	2.024	6.530	65.659	3.378	10.898	58.315
C_5	1.512	4.877	70.536	2.257	7.282	65.597
C_6	1.341	4.326	74.862	1.959	6.318	71.915
C_7	1.015	3.274	78.135	1.928	6.221	78.135

献率普遍得以提高，说明旋转后对主分量的解释更为明确，这7个主分量可以很好的表征原始场大部分信息和主要趋势。相比主分量C_1，其余主分量的方差贡献率都比较小，这说明祁连山及周边的地形影响特征很明显。

以主成分分析得到的高荷载值作为中心点，对荷载值≥0.6各个测站进行聚类评判，当评判结果通过95%显著性水平检验时，可认为分类结果合理，评判标准与结果见表2。通过聚类检验，研究区域构成以乐都、海晏、野牛沟、武威、高台、临泽、刚察为中心点的7个区域(Z1~Z7)，一个区域对应一种类型，从而得到祁连山及周边多个相对独立的分区分型区域。

2.3 聚类结果特征分析

依据聚类检验分析结果，对各个降水区域分布特征进行研究(图2)。结果显示：第一区域Z1位于研究区域东南部，包括乌鞘岭、天祝、永登、乐都、平安、民和、化隆、循化8个站点，占研究区域测站比例为25.8%，最大荷载中心在乐都(0.892)，平均荷载值0.376，年平均降水量296.3 mm，海拔1800~3000 m，属于研究区域季风气候响应区，是祁连山及周边自然生态门户以及农牧业重点发展地区。第二区域Z2位于祁连山中南部，包括门源、海晏、湟源、大通、互助、西宁、湟中7个站点，测站占比为22.6%，最大荷载值0.751，年平均降水量398.5 mm，海拔2500~3000 m，是研究区域草原与牧场基地，也是祁连山区重要的水源涵养地。第三区域Z3分布在沿祁连山中段，包括山丹、民乐、肃南、托勒、野牛沟、祁连6个站点，测站占比为19.3%，最大荷载值0.786，年平均降水量285.0 mm，海拔1700~3400 m，是祁连山重要生态保护与植被恢复地区。第四区域Z4分布在祁连山北麓东端即石羊河流域，测站占比为12.9%，最大荷载值0.903，年平均降水量170.3 mm，海拔1300~2000 m，是发端于冷龙岭北侧大雪山的重要绿洲流域，属于典型的灌溉农业。第五区域Z5为最北端疏勒河流域，测站占比6.5%，最大荷载值0.824，年平均降水量仅有81.2 mm，海拔1400 m，是源于祁连山西段托来南山与疏勒南山之间的绿洲地带，属于灌溉农业。第六区域Z6地处黑河流域，发端于祁连山北麓中段冰川地带，测站占比为6.5%，最大荷载值0.952，年平均降水量100.4 mm，海拔1470 m，属于灌溉农业。第七区域Z7为祁连山海拔最高区域，包括天峻、刚察，测站占比6.5%，最大荷载值0.814，年平均降水量341.3 mm，海拔3400 m，位居大通山地段，是祁连山多条河流的发源地。综上可以看出，分区分型结果与降水量、地理地形、海拔高度显著相关，各个分区相互独立，差异性明显，同一区域特征相近，具有共性，从而构成祁连山独特的降水空间布局特征。

依据聚类检验方法，可将祁连山及周边划分为7个区域，其结论与研究成果[25-26]相吻合：受西风带、副热带、青藏高原环流的共同影响，特别是山地地形的热力和动力作用，祁连山及周边容易构成气旋式环流，副热带高压明显增强，西南暖湿气流强化，夏季风显著北延，冷暖空气一旦相遇，极易在祁连山一带形成强降水区，西北气流型、低槽东移型、东高西低型均是构成祁连山强降水的有利天气形势。也就是说，局地环流、边界层结构均能有效促使祁连山地形正反馈，导致降水量显著增加，且从东南到西北逐渐减少，从而构成独特的降水空间变

表2 祁连山及周边降水量聚类中心以及分区分型结果

Tab. 2 Cluster center site and regional classification of precipitation in Qilian Mountains and its surrounding areas

降水区域	中心站	最大荷载值	站点数量	统计量	临界值
Z1	乐都	0.892	8	34.303	40.113
Z2	海晏	0.751	7	7.607	31.410
Z3	野牛沟	0.786	6	2.054	23.685
Z4	武威	0.903	4	0.501	11.070
Z5	高台	0.824	2	0.000	3.841
Z6	临泽	0.952	2	0.000	3.841
Z7	刚察	0.814	2	0.000	3.841

注：临界值是指在给定信度0.05下具有自由度η-1随机变量χ^2的检验值。

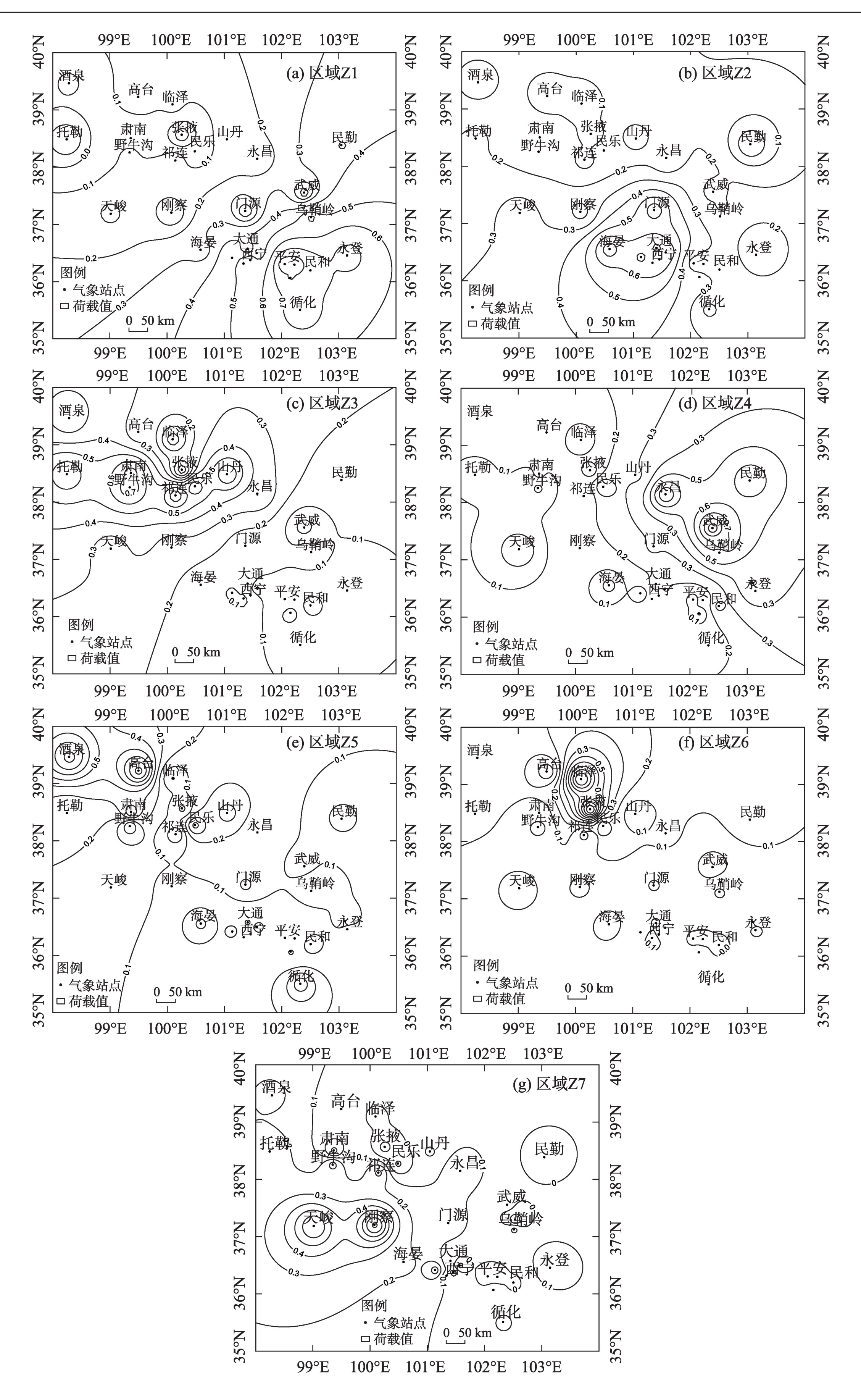


图2 祁连山及周边各个区域荷载值的空间分布

Fig. 2 Spatial distributions of the loading values of each divided areas in Qilian Mountains and its surrounding areas

化特征。

2.4 典型流域增雨效果评价

2.4.1 增雨效果评价背景 统计检验的评估对象是地面降水量，基本方法通过同一时期对比区自然降水量与影响区实测降水量的比较，得出定量化结果，这个结果就是增雨效果。但在作业影响区确定以后，不同对比区就会有不同的增雨效果。聚类检验分区分型结果，可以作为对比区选择的基本依据，换一句话说，当影响区和对比区位于同一区域时，统计检验才有了可信度。

选取Z4，即石羊河流域作为实证区域。该流域人口密集、水资源匮乏、生态环境极其脆弱，人工增雨已成为缓解旱情与生态治理的重要措施。现已建成人工影响作业点71个，其中火箭增雨作业点45个、高炮增雨防雹作业点14个、固定烟炉作业点8个、焰弹作业点4个，覆盖面积6000 km^{2}以上。该流域人工影响作业始于1992年，截至目前，累积作业6000余次，发射高炮25000余发、火箭16000余枚、焰弹20000余枚，具备对比区与影响区遴选条件。

首先，Z4所在区域500~700 hPa常年盛行西北风向，只要对比区位于影响区上风方或者侧风方，就可以使对比区不受催化作业影响而保持相对的独立性；其次，Z4所在区域地处大陆腹地，气候干旱、降水少、蒸发量大，地形地貌、气候背景、植被特征、区域面积、天气系统高度一致，地面降水量显著相关。石羊河流域地理地形以及增雨作业点布局见图3。

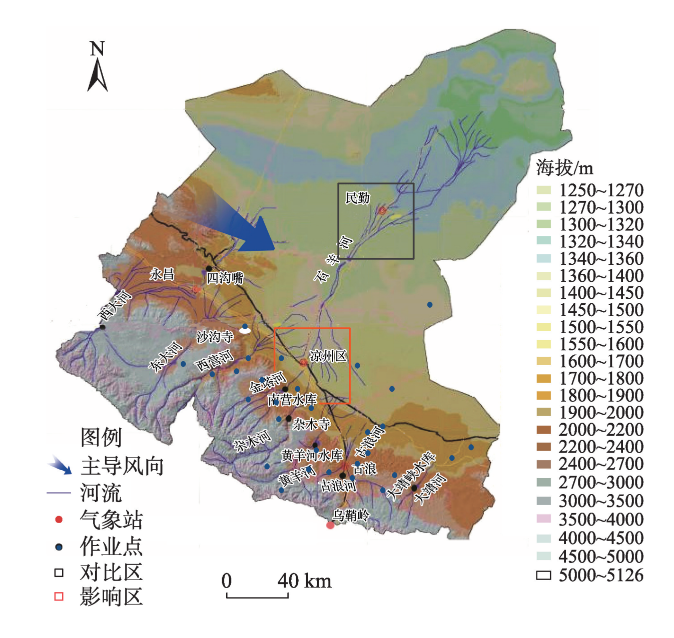


图3 石羊河流域地理分布以及对比区与影响区布局

Fig. 3 Geographical distribution of Shiyang River Basin and the layout of contrast area and affected area

2.4.2 增雨效果评价和检验 石羊河流域地面作业集中在5—9月，选取民勤为对比区，凉州为影响区，即对比区位于影响区侧风方。将1961—1991年作为历史降水量序列时段，1992—2020年作为作业影响序列时段，2个时段样本量基本相当。在应用区域历史回归统计检验方法之前，需要确定统计变量是否满足正态分布，为了使回归模型具有可比性，要求对比区和影响区在2个时段同时具备正态性，即通过Kolmogorov-Smimov（KS）检验[27]。检验与分析结果见表3。

以对比区降水量为自变量，影响区降水量为因变量，运用地面降水量区域历史回归统计检验方法对该时段降水量进行分析。对应时段回归方程均通过95%的显著性水平检验，表明拟合方程可以代表该时段降水量的基本特征（表3）。将对比区降水量代入回归方程，得到影响区降水量的预估值，在

表3 石羊河流域不同时段地面作业效果检验的特征值

Tab. 3 Eigenvalues of ground operation effect test at different periods in Shiyang River Basin

月份	Z(Sig.)		相关系数(r)	回归方程	T	绝对增雨量/mm	相对增雨率/%
	对比区	影响区					
5	1.015(0.254)	0.935(0.346)	0.833*	y=5.929+0.950x	8.110	−0.19	−1.04
6	1.242(0.092)	1.257(0.085)	0.604*	y=10.411+0.972x	4.084	−1.56	−6.26
7	0.864(0.444)	0.905(0.386)	0.490*	y=18.367+0.385x	3.029	6.30	21.86
8	0.948(0.330)	1.051(0.220)	0.680*	y=12.292+0.832x	4.998	5.44	16.11
9	1.261(0.083)	0.877(0.426)	0.677*	y= 9.889+0.849x	4.948	1.26	4.40
5—9	0.502(0.963)	0.564(0.908)	0.665*	y=54.432+0.800x	4.798	8.91	6.51

注：Kolmogorov-Smimov（KS）非参数检验中Z为检验统计量，当双尾渐进概率Sig.≥0.05时，说明样本具有正态性；*表示相关系数通过95%显著性水平检验；T为线性回归方程统计检验值[28]，当T≥0.367时，认为回归关系通过95%显著性水平检验；绝对增雨量=影响区降水量－对比区降水量；相对增雨率=(影响区降水量/对比区降水量－1)×100%。

此基础上对不同时段作业效果展开分析。

1992—2020年5—9月绝对增雨量为8.91 mm，相对增雨率为6.51%，说明5—9月作业效果明显，具有很好的增雨潜力。但在不同月份作业效果却表现出明显的差异性(表3)。5—6月为负效果，7—9月为正效果，其中7月最高(6.30 mm，21.86%)，8月次之(5.44 mm，16.11%)，这一结果与石羊河流域气候背景以及天气系统演变规律密切关联。5—6月正值春末夏初，是典型的“卡脖子”时段，从多年降水量分布上看，该时段降水量不足全年的20%，水汽条件差、增雨效率不高、作业效果弱，实际业务中应该慎重选择作业时机，尽可能降低负效果的出现几率。而7—9月正是高原降水云系活跃阶段，尤其是7—8月，阵性降水多、水汽条件好、增雨效率高、作业效果明显，可以通过增加作业点次以及作业量，显著增加降水量，有效提升作业效果。

3 讨论

作业效果评价一直是制约人工影响天气发展的技术性难题。祁连山地形复杂、范围广大，我们前期已经对降水分布的区域性特征进行了一些研究，在利用统计检验评价增雨效果时，发现在对比区的选择上存在诸多问题，导致检验效果出现很多不确定性结果。众所周知，对比区与影响区降水量比对是统计检验的基本思路，大量事实证明，在作业影响区确定以后，不同对比区就会有不同的增雨效果。遗憾的是，一直以来，关于对比区选择并未形成客观准确的方法，也没有可以依靠的理论依据，地面作业效果评价与人工增雨服务陷入困境。

聚类检验分区分型结果，可以作为对比区选择的基本依据。只有当影响区和对比区位于同一区域时，统计检验才具有可靠的前提条件。本文采用主成分分析方法对研究区域降水场进行分区分型，再通过聚类检验进行结果验证，很大程度上保证了分区的准确性。将分区分型结果应用于对比区与影响区的选择，解决了在增雨效果评价上区域选择的盲目性问题，这也是制约地面增雨效果评价的关键环节。这种方法通常用于气候分型区划，但很少用于区域降水研究，在人工增雨效果检验对比区选择上则更为少见。祁连山以及周边区域地形复杂，降水分布极为不均，建立科学准确的分区方法，对把握该区域降水特征具有极为重要的意义。

4 结论

基于主成分分析与聚类分析的一致性，本文对分区分型方法进行了深入研究，并验证了这种结果的合理性。通过分析，将祁连山及周边降水分布划分为7个区域，在对上述区域逐一评价的基础上，选择石羊河流域作为实例。设计对比区、影响区，采用区域历史回归统计方法对不同时段降水量展开增雨效果评价。结果显示：1992—2020年5—9月绝对增雨量8.91 mm，相对增雨率6.51%，其中7月最高(6.30 mm，21.86%)，8月次之(5.44 mm，16.11%)，说明人工增雨效果在不同时段具有差异性，适时作业、科学作业，才能提高综合效果，最大程度提升抗旱蓄水、生态治理与修复的能力和水平。

参考文献(References)

[1] 杨昭明, 张调风. 1961—2017年青藏高原东北部雨季降水量变化及其贡献度分析[J]. 干旱区研究, 2021, 38(1): 22-28. [Yang Zhaoming, Zhang Tiaofeng. Analysis of precipitation change and its contribution in rainy season in northeast Qinghai Tibet Plateau from 1961 to 2017[J]. Arid Zone Research, 2021, 38(1): 22-28.]

[2] 王春林, 司建华, 赵春彦, 等. 河西走廊近57年来干旱灾害特征时空演化分析[J]. 高原气象, 2019, 38(1): 196-205. [Wang Chunlin, Si Jianhua, Zhao Chunyan, et al. Analysis of temporal and spatial evolution of characteristics of drought disaster in the Hexi Corridor in recent 57 years[J]. Plateau Meteorology, 2019, 38(1): 196-205.]

[3] 陈乾, 陈添宇, 肖宏斌. 祁连山区夏季各类降水过程的典型个例分析[J]. 干旱气象, 2008, 26(3): 1-7. [Chen Qian, Chen Tianyu, Xiao Hongbin. Typical cases study of classified precipitation process in summer over Qilian Mountains[J]. Arid Meteorology, 2008, 26(3): 1-7.]

[4] 郭小芹, 罗永忠. 1961—2016年祁连山区季节性干旱综合指数特征分析[J]. 气象与环境学报, 2019, 35(1): 88-93. [Guo Xiaoqin, Luo Yongzhong. Study on characteristics of seasonal drought composite index in Qilian Mountain area from 1961 to 2016[J]. Journal of Meteorology and Environment, 2019, 35(1): 88-93.]

[5] 曾光平, 方仕珍, 肖锋. 1975—1986年古田水库人工降雨效果总分析[J]. 大气科学, 1991, 15(4): 97-108. [Zeng Guangping, Fang Shizhen, Xiao Feng. The total analysis of the effect of artificial rainfall in Gutian reservoir area, Fujian (1975—1986)[J]. Chinese Journal of Atmospheric Sciences, 1991, 15(4): 97-108.]

[6] 王婉, 姚展予. 2006年北京市人工增雨作业效果统计分析[J]. 高原气象, 2009, 28(1): 195-202. [Wang Wan, Yao Zhanyu. Sta-

tistical estimation of artificial precipitation enhancement effectiveness in Beijing in 2006[J]. Plateau Meteorology, 2009, 28(1): 195-202.]

[7] 常倬林, 崔洋, 张武, 等. 基于CERES的宁夏空中云水资源特征及其增雨潜力研究[J]. 干旱区地理, 2015, 38(6): 1112-1120. [Chang Zhuolin, Cui Yang, Zhang Wu, et al. Characteristics and developing potential of cloud water resource in Ningxia with the CERES data[J]. Arid Land Geography, 2015, 38(6): 1112-1120.]

[8] 严采蘩, 陈万奎. 统计评价人工增雨效果的一些问题研究[J]. 应用气象学报, 2001, 12(增刊1): 80-86. [Yan Caifan, Chen Wankui. Issues existed in the statistical assessment of artificial rainfall effect[J]. Journal of Applied Meteorology, 2001, 12(Suppl. 1): 80-86.]

[9] 许焕斌. 关于在人工影响天气中更新学术观念的探讨[J]. 干旱气象, 2009, 27(4): 305-307. [Xu Huanbin. On the ideal renew in weather modification[J]. Journal of Arid Meteorology, 2009, 27(4): 305-307.]

[10] 孙倩倩, 刘晶淼. 基于聚类分析的中国东北地区气温和降水时空变化特征[J]. 气象与环境学报, 2014, 30(3): 59-65. [Sun Qianqian, Liu Jingmiao. Temporal-spatial change of temperature and precipitation based on CAST in northeast China[J]. Journal of Meteorology and Environment, 2014, 30(3): 59-65.]

[11] 韩微, 翟盘茂. 三种聚类分析方法在中国温度区划分中的应用研究[J]. 气候与环境研究, 2015, 20(1): 111-118. [Han Wei, Zhai Panmao. Three cluster methods in regionalization of temperature zones in China[J]. Climatic and Environmental Research, 2015, 20(1): 111-118.]

[12] 么枕生. 聚类统计检验用于气候分类[M]. 北京: 气象出版社, 1997: 85-94. [Yao Zhensheng. Cluster statistical test for climate classification[M]. Beijing: Meteorological Press, 1997: 85-94.]

[13] 邱美娟, 宋迎波, 王建林, 等. 新型统计检验聚类方法在精细化农业气象产量预报中的应用[J]. 中国农业气象, 2014, 35(2): 187-194. [Qiu Meijuan, Song Yingbo, Wang Jianlin, et al. Application of newly developed cluster analysis of statistical test in fine agrometeorological yield prediction[J]. China Journal of Agrometeorology, 2014, 35(2): 187-194.]

[14] 全占军, 李远, 李俊生, 等. 采煤矿区的生态脆弱性——以内蒙古锡林郭勒草原胜利煤田为例[J]. 应用生态学报, 2013, 24(6): 1729-1738. [Quan Zhanjun, Li Yuan, Li Junsheng, et al. Ecological vulnerability of coal mining area: A case study of Shengli Coalfield in Xilinguole of Inner Mongolia, China[J]. Chinese Journal of Applied Ecology, 2013, 24(6): 1729-1738.]

[15] 李湘阁, 胡凝. 实用气象统计方法[M]. 北京: 气象出版社, 2015: 11-28. [Li Xiangge, Hu Ning. Practical meteorological statistical method[M]. Beijing: Meteorological Press, 2015: 11-28.]

[16] 丁裕国, 张耀存, 刘吉峰. 一种新的气候分型区划方法[J]. 大气科学, 2007, 31(1): 129-136. [Ding Yuguo, Zhang Yaocun, Liu Jifeng. A new method of climate classification and compartment using the conjunction between CAST and REOF[J]. Chinese Journal of Atmospheric Sciences, 2007, 31(1): 129-136.]

[17] 刘吉峰, 李世杰, 丁裕国. 等. 一种用于中国年最高(低)气温区划的新的聚类方法[J]. 高原气象, 2005, 24(6): 966-973. [Liu Jifeng, Li Shijie, Ding Yuguo, et al. A new cluster method applied to yearly highest and lowest air temperature region division in China [J]. Plateau Meteorology, 2005, 24(6): 966-973.]

[18] 谢飞, 顾继光, 林彰文. 基于主成分分析和熵权的水库生态系统健康评价——以海南省万宁水库为例[J]. 应用生态学报. 2014, 25(6): 1773-1779. [Xie Fei, Gu Jiguang, Lin Zhangwen. Assessment of aquatic ecosystem health based on principal component analysis with entropy weight: A case study of Wanning Reservoir[J]. Chinese Journal of Applied Ecology, 2014, 25(6): 1773-1779.]

[19] 孙莉娟, 陈金华, 徐阳, 等. 基于多指数安徽省强降水灾害近45年的时空演变[J]. 水土保持研究, 2020, 27(3): 128-133. [Sun Lijuan, Chen Jinhua, Xu Yang, et al. Spatial and temporal evolution of heavy rainfall disasters in Anhui Province in recent 45 years based on multi-indictors[J]. Research of Soil and Water Conservation, 2020, 27(3): 128-133.]

[20] 韩群柱, 冯起, 高海东, 等. 基于主成分分析的关中地区农业粮食生产变化的影响因素研究[J]. 干旱区地理, 2020, 43(2): 474-480. [Han Qunzhu, Feng Qi, Gao Haidong, et al. Influencing factors of agricultural production changes in Guanzhong based on PCA[J]. Arid Land Geography, 2020, 43(2): 474-480.]

[21] 韩微, 翟盘茂. 三种聚类分析方法在中国温度区划分中的应用研究[J]. 气候与环境研究, 2015, 20(1): 111-118. [Han Wei, Zhai Panmao. Three cluster methods in regionalization of temperature zones in China[J]. Climatic and Environmental Research, 2015, 20(1): 111-118.]

[22] 冯克鹏, 田俊灿, 沈辉. 基于K-means聚类分区的西北地区近半个世纪气温变化特征分析[J]. 干旱区地理, 2019, 42(6): 1239-1252. [Feng Kepeng, Tian Juncan, Shen Hui. Temperature variation characteristics of northwest China based on K-means clustering partition in the past half century[J]. Arid Land Geography, 2019, 42(6): 1239-1252.]

[23] 雷月, 宫彦龙, 邓茹月, 等. 基于主成分分析和聚类分析综合评价蒸谷米的品质特性[J]. 食品工业科技, 2021, 42(7): 258-267. [Lei Yue, Gong Yanlong, Deng Ruyue, et al. Comprehensive evaluation of quality characteristics of parboiled rice based on principal component analysis and cluster analysis[J]. Science and Technology of Food Industry, 2021, 42(7): 258-267.]

[24] 邓北胜. 人工影响天气技术与管理[M]. 北京: 气象出版社, 2011: 140-147. [Deng Beisheng. Technology and management of weather modification[M]. Beijing: China Meteorological Press, 2011: 140-147.]

[25] 梁潇云, 刘屹岷, 吴国雄. 热带、副热带海陆分布与青藏高原在亚洲夏季风形成中的作用[J]. 地球物理学报, 2006, 49(4): 983-992. [Liang Xiaoyun, Liu Yimin, Wu Guoxiong. Roles of tropical and subtropical land-sea land distribution and the Qinghai-Xizang

Plateau in the formation of the Asian summer monsoon, China[J]. Chinese Journal of Geophysics, 2006, 49(4): 983-992.]

[26] 高荣, 韦志刚, 钟海玲. 青藏高原陆表特征与中国夏季降水的关系研究[J]. 冰川冻土, 2017, 39(4): 741-747. [Gao Rong, Wei Zhigang, Zhong Hailing. Relationship between land surface characteristics in the Tibetan Plateau and summer precipitation in China[J]. Journal of Glaciology and Geocrylogy, 2017, 39(4): 741-747.]

[27] 黄嘉佑. 气象统计分析与预报方法[M]. 第三版. 北京: 气象出版社, 2004. [Huang Jiayou. Meteorological statistical analysis and forecasting methods[M]. 3rd ed. Beijing: Meteorological Press, 2004.]

[28] 张敏静, 刘雅娜, 薛志群. 一元线性回归方程有关检验问题的研究[J]. 价值工程, 2012, 31(2): 1-2. [Zhang Minjing, Liu Yana, Xue Zhiqun. Linear regression equation study of relevant inspection question[J]. Value Engineering, 2012, 31(2): 1-2.]

Cluster analysis with statistical test of precipitation distribution in Qilian Mountains and its surrounding area and evaluation of artificial precipitation enhancement in typical watershed

GUO Xiaoqin[1,2], LI Guangming[2], SUN Zhanfeng[2], WANG Xingtao[2]

(1. Key Laboratory for Cloud Physical Environment of China Meteorological Administration, Beijing 100081, China;
2. Wuwei Meteorological Bureau, Wuwei 733000, Gansu, China)

Abstract: The topography of Qilian Mountains and its surrounding areas in northwest China is complex, and the precipitation it receives is extremely irregular in time and space. The ability to characterize the distribution of rainfall and snowfall has great practical significance for understanding changes in precipitation patterns. Using data from 31 base stations covering the months of May-September, 1961—2020, precipitation was analyzed by principal component analysis (PCA). The significance of the results was then tested by cluster analysis with statistical test (CAST). Finally, a geographic zonation was applied to evaluate the effects of artificial precipitation enhancement. The results show that: (1) Seven geographic areas (Z1-Z7) can satisfactorily represent the characteristics of precipitation distribution in the study area (the contribution rate of cumulative variance exceeds 78%). (2) The zones take Ledu, Haiyan, Yeniugou, Wuwei, Gaotai, Linze, and Gangcha stations as the central points respectively, and their zonal characteristics significantly reflect precipitation, terrain, and altitude. (3) Effects of artificial precipitation enhancement were evaluated by statistical regression of regional historical data. For the months of May-September during 1992—2020, the benefit was most pronounced in the Shiyang River Basin, where the absolute enhancement was 8.91 mm and the relative precipitation enhancement rate was 6.51% with the highest in July (6.30 mm, 21.86%), followed by August (5.44 mm, 16.11%). Assessing artificial enhancement of precipitation based on ground data relies heavily on the selection of a suitable comparison area, so results can change significantly with different comparison choices. The method of cluster analysis with statistical testing is not only helpful in zoning and classification of precipitation in complex terrain, but it also provides an objective basis for the selection of comparison areas.

Key words: cluster analysis with statistical test (CAST); principal component analysis (PCA); division and classification; evaluation of precipitation enhancement effects